

PATENT APPLICATION  
OUR REF. NO.: SAM2.0002

METHOD OF DETECTING MOTION IN AN INTERLACED VIDEO SEQUENCE  
BASED ON LOGICAL OPERATION ON LINEARLY SCALED MOTION  
INFORMATION AND MOTION DETECTION APPARATUS

By:

Yeong-Taeg Kim

and

Shin-Haeng Kim

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

I hereby certify that this paper of fee is being deposited with the United States Postal Service on this date: 10-22-01,  
in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number FESDE2436384US addressed to:  
Box Patent Application Assistant Commissioner for Patents, Washington, D.C. 20231

  
Print Name: Michelle Daugherty

METHOD OF DETECTING MOTION IN AN INTERLACED VIDEO SEQUENCE  
BASED ON LOGICAL OPERATION ON LINEARLY SCALED MOTION  
INFORMATION AND MOTION DETECTION APPARATUS

5

*Ins Q17*

BACKGROUND OF THE INVENTION

CROSS-REFERENCES TO RELATED APPLICATIONS

Applicants claim the benefit of U.S. Provisional Application No. 60/257,338 entitled  
10 "Methods of Detecting Motion in an Interlaced Video Sequence Based on Logical  
Operation on the Linearly Scaled Motion Information and the Apparatus Therefor," filed  
December 20, 2000, U.S. Provisional Application No. 60/257,365 entitled "Methods of  
Detecting Motion in an Interlaced Video Sequence Utilizing Region-Wise Motion and  
Apparatus" filed December 20, 2000, and U.S. Provisional Application No. 60/273,100  
15 entitled "Method of Detecting Repetitive Motion In An Interlaced Video Sequence and  
Apparatus Therefore," filed March 2, 2001, which applications are incorporated herein by  
reference.

10001676-102301

FIELD OF THE INVENTION

20 The present invention lies in the signal processing field. More specifically, the  
present invention pertains to a method of detecting motion in an interlaced video sequence.  
The invention is particularly applicable to the conversion of an interlaced video signal to a  
progressive frame video signal, whereby regional motion information is utilized to define  
whether or not the video sequence contains motion or represents still image information.  
25 The present invention also pertains to an apparatus for performing the claimed method.

## DESCRIPTION OF THE RELATED ART

In the development of current digital TV (DTV) systems, it is essential to employ a video format conversion unit because of the variety of the video formats adopted in many different DTV standards worldwide. For instance, the ATSC DTV standard system of the

5 North America adopted 1080x1920 interlaced video, 720x1280 progressive video, 720x480 interlaced and progressive video, and so on, as its standard video formats for digital TV broadcasting. Video format conversion refers to a signal processing operation in which an incoming video format is converted to a specified output video format so that the output video can be properly displayed on a displaying device such as a monitor, FLCD, or a

10 Plasma display, which has a fixed resolution.

Video format conversion systems are of significant importance since the conversion can directly affect the visual quality of the video of a DTV receiver. Fundamentally, the video format conversion operation requires advanced algorithms for multi-rate system design, poly-phase filter design, and interlaced-to-progressive scanning rate conversion or simply deinterlacing. Deinterlacing represents an operation that doubles the vertical scanning rate of the interlaced video signal.

Interlaced video in general is a sequence of separately arriving fields, such as A1,

20 A2, A3, etc., where A1, A2, and A3 are interlaced images with A1 being a top image, A2 being a bottom image, A3 being the next top image, and so on. The most popular systems currently in use, namely NTSC, PAL, and SECAM are two-field systems, where two consecutive fields (such as the top field A1 and the bottom field A2) make up a frame. Each scanned field contains, i.e., updates, every other line of a corresponding frame and the

25 number of lines in the frame is twice the number of lines in each of the fields which contain

video information. Typically, the first field is identified with odd-numbered lines and the second field is identified with even-numbered lines. The fields are scanned onto the display screen one after the other at a defined frequency.

5 By way of example, NTSC scans close to 30 frames (60 fields of interlaced video) per second, with 525 lines per frame, and a horizontal to vertical aspect ratio of 4:3. The frame difference, therefore, is the difference between two fields having the same types (top or bottom) such as A1 and A3, or A2 and A4. PAL and SECAM scan 25 frames per second, with 625 lines per image, and the same aspect ratio of 4:3. As noted, the interlacing in all of  
10 these systems is 2:1, i.e., two fields per one frame. The primary reason for the interlacing of the lines between the fields is to reduce flicker in the display. An image that is updated, say, only 30 times a second would allow the human eye to perceive the scanning, because the image information would already start to fade before the next image is scanned onto the screen. When two fields are used, and each contains half of the information, the scanning  
15 rate in effect is raised to 60 Hz, and the human eye no longer perceives any flicker.

Deinterlacing refers to the filling of unavailable lines in each of the fields A1, A2, A3, and so on. As a result of deinterlacing, a 60 Hz field sequence (of interlaced video fields) becomes a 60 Hz progressive sequence.

20 Interlaced video is subject to several intrinsic drawbacks, referred to as artifacts. These include serrated lines that are observed when there is motion between fields, line flickering, raster line visibility, and field flickering. These also apply to DTV (digital TV) receivers. Historically, deinterlacing algorithms have been developed to enhance the video  
25 quality of NTSC TV receivers by reducing these intrinsic annoying artifacts of the

interlaced video signal. Besides, elaborate deinterlacing algorithms utilizing motion detection or motion compensation provide excellent methods of doubling the vertical scanning rate of the interlaced video signal especially for stationary (motionless) objects in the video signal.

5

The present invention therefore also relates to the motion detection based deinterlacing operation that can be used for analog and digital TV receivers.

The state of the art includes a variety of deinterlacing algorithms, each having been  
10 exploited and studied comprehensively by many researchers during the last decade.

Deinterlacing algorithms can be categorized into two classes, namely, 2-D (spatial) deinterlacing algorithms and 3-D (spatio-temporal) deinterlacing algorithms depending on the use of motion information embedded in consecutive interlaced video sequence.

Combined spatial and temporal 3-D deinterlacing algorithms based on a motion detection give more pleasing performance than 2-D deinterlacing algorithms. The key point of a 3-D deinterlacing algorithm is how to precisely detect motion in the interlaced video signals.

The publications in the following list disclose some of the applicable deinterlacing methods. They may be categorized as follows:

20 [1] Simple line doubling scheme, vertical filtering, vertical edge controlled interpolation method disclosed in the IEEE Transactions on Consumer Electronics, pp. 279-89, August 1989 by D.I. Hentschel;

[2] Edge direction dependent deinterlacing method disclosed in the Proc. of the Int. Workshop on HDTV, 1994, by D. Bagni, R Lancini, and S. Tubaro;

10001676-102227

[3] Nonlinear interpolation methods based on:

a weighted median filter disclosed in the Proc. of the IEEE ISCAS, pp. 433-

36, Portland, USA, May 1989, by J. Juhola, A. Nieminen, J. Sal, and Y. Neuvo,

5

FIR median hybrid interpolation disclosed in Pro. Of SPIE's Visual Communications and Image Processing, Lausanne, Switzerland, October 1990, 00. 125-32 by A. Lehtonen and M. Renfors,

10

a complementary median filter disclosed in Proc. of the Int. Workshop on HDTV, 1994 by H. Blume, I. Schwoerer, and K. Zygis,

15

[4] A motion adaptive method disclosed in IEEE Transactions on Consumer Electronics, pp. 110-114, May 1990 by C. Markhauser.

More recently, a new motion detection based deinterlacing method has been described in the following two, commonly assigned patents:

20

[5] U.S. Patent No. 5,943,099, Aug. 24, 1999, issued to Young-Taeg Kim, entitled Interlaced-to-Progressive Conversion Apparatus and Method Using Motion and Spatial Correlation. There, an interlaced-to-progressive conversion device includes a spatial interpolator that provides for spatial interpolation and a temporal interpolator that provides for temporal interpolation of an interlaced video input signal. The system reduces jitter and related artifacts by temporally or spatially interpolating the signals.

100016772000

[6] U.S. Patent No. 5,959,681, Sep. 28, 1999, to Yong-Hun Cho, entitled Motion Picture Detecting Method. There, two separate field memories are utilized for detecting rapid motion and slow motion in an interlaced video sequence. An interlaced video signal is thereby converted into a progressive-scanned signal.

5 Differences between spatial interpolation and temporal interpolation are used to determine whether the image is in motion. If the differences exceed certain defined thresholds, motion is determined. The thresholds are dynamically adapted during the process.

10 The core of the methods described in the latter two patents is to estimate a motion decision factor based on the frame difference signal and the sample correlation in the vertical direction. These methods provide a way of reducing the visual artifacts that can be possibly arising from false motion detection by utilizing the sample correlation in vertical direction of the sampling point where the value is to be interpolated. A common drawback  
15 of those methods, however, is that they do not provide a true motion detection method when there are high frequency components in the vertical direction. In other words, when there are high frequency components in the vertical direction, the methods described in the references [5] and [6] will come to the conclusion that motion pictures are processed.

20 As a consequence, in many instances, those prior art processing methods do not provide for an increase in the vertical resolution even when no real motion is present between fields.

## SUMMARY OF THE INVENTION

It is accordingly an object of the invention to provide a motion detection method in an interlaced video sequence, which overcomes the above-mentioned disadvantages of the heretofore-known devices and methods of this general type and which provides for a robust  
5 method of estimating a motion decision parameter which is associated with the point to point degree of motion in the interlaced video sequence. It is another object of the present invention to disclose a deinterlacing method and apparatus by utilizing the motion decision parameter of the invention.

10 With the foregoing and other objects in view there is provided, in accordance with the invention, a method of computing a motion decision value for further utilization in a video signal processing system. The method comprises the following steps:

inputting a video signal with an interlaced video sequence;  
computing a frame difference signal from a difference between a previous field and  
15 a next field in the video sequence;

forming a point-wise motion detection signal from the frame difference signal;  
computing a region-wise motion detection signal from the point-wise motion detection signal and an adjacent point-wise motion detection signal delayed by one field;  
and

20 forming from the region-wise motion detection signal a motion decision value and outputting the motion decision value for further processing in the video signal processing system.

10001626402400

In accordance with an added feature of the invention, the difference signal is low-pass filtered prior to the step of forming the point-wise motion detection signal.

The low-pass filter may be defined by a low pass filter matrix

5

$$W_{M \times N} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \cdots & w_{MN} \end{bmatrix}$$

where  $w_{11}, \dots, w_{MN}$  represent a set of predetermined coefficients.

10 In accordance with an additional feature of the invention, the step of forming the point-wise motion detection signal comprises computing

$$f_n(i, h) = l_K(d_n(i, h))$$

15 where  $f_n$  is the point-wise motion detection signal,  $i$  and  $h$  define a spatial location of the respective video signal value in a cartesian matrix,  $l_K(\cdot)$  denotes a linearly scaling function represented as

$$l_K(y) = \begin{cases} 1, & \text{if } y \geq K \\ y / K, & \text{otherwise} \end{cases}$$

20

in which K is a positive constant value, or, alternatively, with the linearly scaling function

$$l_K(y) = \begin{cases} 1, & \text{if } y \geq K \\ (y/K)^a, & \text{otherwise} \end{cases}$$

5

where K is still constant and  $a$  is a real number.

Various options are available for computing the region-wise motion detection signal from the point-wise motion detection signal. The following equations can be selected

10

$$\phi_n(i, h) = f_n(i, h) + \min(f_{n-1}(i-1, h), f_{n-1}(i+1, h))$$

$$\phi_n(i, h) = \text{med}(f_n(i, h), f_{n-1}(i-1, h), f_{n-1}(i+1, h)), \text{ or}$$

$$\phi_n(i, h) = \max(f_n(i, h), f_{n-1}(i-1, h), f_{n-1}(i+1, h))$$

15

where  $f_{n-1}(\cdot)$  denotes a motion detection signal delayed by one field,  $\text{med}(\cdot)$  denotes a median operation,  $\max(\cdot)$  denotes an operation to minimize an error from a false motion detection, and the indices  $i$  and  $h$  define a spatial location of the respective video signal value in a cartesian matrix.

20

In accordance with another feature of the invention, the region-wise motion detection signal is low-pass filtered prior to the outputting step. In a preferred embodiment, the region-wise motion detection signal is low-pass filtered to form the motion decision value  $m_n(i, h)$  by:

$$m_n(i, h) = \sum_{p=-a}^b \sum_{q=-c}^d \phi_n(i + 2 \times p, h + 2 \times q) \cdot \alpha_{p,q}$$

where  $a, b, c, d \geq 0$ , and  $\alpha_{p,q}$  represents a set of normalized predetermined coefficients of a low pass filter.

5 The kernel of the low-pass filter may be defined as

$$[\alpha_{p,q}'s] = \begin{bmatrix} 0 & 1/8 & 0 \\ 1/8 & 4/8 & 1/8 \\ 0 & 1/8 & 0 \end{bmatrix}.$$

With the above objects in view there is also provided, in accordance with the  
 10 invention, a method of processing interlaced video signals, which comprises:

- spatially interpolating a value of the video signal at a given location from a video signal of at least one adjacent location in a given video field;
- temporally interpolating the value of the video signal at the given location from a video signal at the same location in temporally adjacent video fields; and
- 15 forming a motion decision value for the same location in accordance with the above-outlined method; and
- mixing an output signal for the video signal at the given location from the spatially interpolated signal and the temporally interpolated signal and weighting the output signal in accordance with the motion decision value.

20

1000157-10540

In accordance with again an added feature of the invention, the motion decision value is varied between 0 and 1 as a function of an estimate of the degree of motion at the given location and, upon estimating a high degree of motion, heavily weighting the output signal towards the spatially interpolated signal and, upon estimating a low degree of motion,  
5 heavily weighting the output signal towards the temporally interpolated signal.

In accordance with again an additional feature of the invention, the spatially interpolated signal is output as the output signal upon estimating a high degree of motion, and the temporally interpolated signal is output as the output signal upon estimating a low  
10 degree of motion.

With the above and other objects in view there is also provided, in accordance with the invention, an apparatus for computing a motion decision value in a video processing system. The apparatus comprises:

15 an input for receiving a video signal with an interlaced video sequence;

difference forming means connected to the input for computing a frame difference signal from a difference between a previous field and a next field of a current field to be deinterlaced;

means for forming a point-wise motion detection signal from the frame difference signal, and for computing a region-wise motion detection signal from the point-wise motion detection signal and an adjacent point-wise motion detection signal delayed by one field;  
20 and

means for forming from the region-wise motion detection signal a motion decision value and for outputting the motion decision value for further processing in the video signal processing system.

5       The apparatus is configured and/or programmed to perform the method outlined above.

A corresponding apparatus of processing interlaced video signals, for example for interlaced to progressive conversion, comprises:

10       an input for receiving a video signal with an interlaced video sequence of fields;

            a spatial interpolator connected to the input and configured to spatially interpolate a value of the video signal at a given location from a video signal of at least one adjacent location in a given video field;

            a temporal interpolator connected to the input in parallel with the spatial interpolator

15       for temporally interpolating the value of the video signal at the given location from a video signal at the same location in temporally adjacent video fields; and

            a computing apparatus as outlined above connected to the input and in parallel with the spatial interpolator and the temporal interpolator for forming a motion decision value for the same location; and

20       a mixer connected to receive an output signal from each of the spatial interpolator, the temporal interpolator, and the computing apparatus, the mixer being configured to mix an output signal for the video signal at the given location from the spatially interpolated

signal and the temporally interpolated signal in dependence on the motion decision value output by the computing apparatus.

Other features which are considered as characteristic for the invention are set forth  
5 in the appended claims.

Although the invention is illustrated and described herein as embodied in a method of detecting motion in an interlaced video sequence based on logical operation on the linearly scaled motion information and an apparatus therefor, it is nevertheless not intended  
10 to be limited to the details shown, since various modifications and structural changes may be made therein without departing from the spirit of the invention and within the scope and range of equivalents of the claims.

15 The construction of the invention, however, together with additional objects and advantages thereof will be best understood from the following description of the specific embodiment when read in connection with the accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWING

Fig. 1 is a diagrammatic view of two juxtaposed fields of a frame of an interlaced  
20 video sequence;

Fig. 2 is a diagrammatic illustration of three fields serving to describe the deinterlacing problem;

25 Fig. 3 is a more detailed view illustrating the deinterlacing process;

Fig. 4 – 6 are block diagrams illustrating the computation of a motion decision parameter according to three algorithms; and

Fig. 7 – 9 are block diagrams illustrating the computation of the motion decision 5 parameter, as illustrated in Figs. 4-6, and the resultant mixing of the spatially and temporally interpolated signals in dependence on the motion decision.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring now to the figures of the drawing in detail and first, particularly, to the 10 introductory view of Fig. 1, an interlaced video sequence is a series of frames each including a plurality of fields. As noted above, all conventional systems utilize two fields per frame which are sequentially scanned. A top field 1 contains information regarding the first ( $v = 0$ ), third ( $v = 2$ ), fifth ( $v = 4$ ), etc. lines, and a bottom field 2 contains information regarding the second ( $v = 1$ ), fourth ( $v = 3$ ), sixth ( $v = 5$ ), etc. lines.

15

In order to systematically describe the deinterlacing problem and the methods of the present invention, let  $x_n$  denote the incoming interlaced video field at a time instant  $t = n$  and  $x_n(v, h)$  denote the associated value of the video signal at the geometrical location  $(v, h)$ .

The variable  $v$  represents the vertical location and  $h$  represents horizontal location, in the 20 cartesian matrix system commonly applied. By definition, the signal values of  $x_n$  of the interlaced video signal are available only for the even lines ( $v = 0, 2, 4, \dots$ ) if  $x_n$  is the top field

1. Similarly, the signal values of  $x_n$  are available only for the odd lines of  $v$  ( $v = 1, 3, 5, \dots$ ) if  $x_n$  is the bottom field 2. Conversely, the signal values of  $x_n$  are not available for odd lines if 25  $x_n$  is a top field signal and the signal values of  $x_n$  are not available for even lines if  $x_n$  a bottom field. Fig. 1 shows the top field 1 scanned at  $t = m$  and the bottom field 2 scanned at

10001657492004

$t = m + 1$ . Top and bottom fields are typically available in turn in the time axis, that is, the top and bottom fields are sequentially scanned to make up a frame.

Based upon the description of the interlaced video signal, deinterlacing problem can

5 be stated as a process to reconstruct or interpolate the non-available signal values of each field. That is, the deinterlacing problem is to reconstruct the signal values of  $x_n$  at the odd lines ( $v = 1, 3, 5, \dots$ ) for the top field  $x_n$  and to reconstruct the signal values of  $x_n$  at the even lines ( $v = 0, 2, 4, \dots$ ) for the bottom field  $x_n$ .

10 For the simple description of the present invention, and to avoid any notational confusion in the disclosure, the deinterlacing problem will be simplified as a process which reconstructs or interpolates the unavailable signal value of  $x_n$  at the  $i^{\text{th}}$  line where the signal values of the lines at  $i \pm 1, i \pm 3, i \pm 5, \dots$  are available. More simply deinterlacing is to  
15 interpolate the value of  $x_n(i, h)$ , which is not originally available. It must be noted that, since  $x_{n-1}$  and  $x_{n+1}$  have a different sampling phase from  $x_n$ , the signal values of  $x_{n-1}(i, h)$  and  $x_{n+1}(i, h)$  are available, which is why motion detection can be incorporated with the deinterlacing problem. This relationship is depicted in Fig. 2, where dotted lines (and white circles) represent “no data available” and solid lines (and black circles) represent “available data.”

20

The deinterlacing method and an exemplary apparatus of the present invention with preferred embodiments will be better understood from the following description, which will make specific reference to Figs. 3-5 of the drawing.

Referring now to Fig. 4, there is illustrated the novel method of estimating a motion decision parameter  $m_n(i,h)$ . Fundamentally,  $m_n(i,h)$  is estimated from the incoming interlaced video sequence and associated with the point-to-point degree of motion in the interlaced video sequence. The importance or the usefulness of estimating  $m_n(i,h)$  can be

5 easily understood from Figs. 2 and 3. Suppose that precise motion detection information is available when we interpolate  $x_n(i,h)$  and suppose there is no motion at the spatial location  $(i,h)$ , then the best interpolation for  $x_n(i,h)$  is to use the value of  $x_{n-1}(i,h)$ . This follows logically from the fact that no motion is introduced between  $t = n - 1$  and  $t = n + 1$  at the spatial location  $(i,h)$ , which very strongly implies that the value of  $x_n(i,h)$  would be close to

10 the value of  $x_{n-1}(i,h)$ . The usage of the motion decision parameter of the present invention is also to utilize the motion information for deinterlacing to properly mix the temporal information, which will be described later.

First, the frame difference signal  $D_n$  is computed by taking the difference between

15 the fields in one frame interval as

$$D_n = |x_{n+1} - x_{n-1}|$$

which associates with the scene change that occurred between the fields  $x_{n+1}$  and  $x_n$ .

20 1. The frame difference signal is then low pass filtered to form

$$d_n = LPF(D_n)$$

where  $LPF(\cdot)$  represents a low pass filtering process over the input video signal. The

25 M x N kernel,  $W_{MxN}$ , in general, of the low pass filter  $LPF(\cdot)$ , can be expressed as

$$W_{M \times N} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \cdots & w_{MN} \end{bmatrix}$$

where  $(w_{11}, \dots, w_{MN})$  represents a set of predetermined coefficients. It should be mentioned that the characteristic of the LPF  $(\cdot)$  can be all-pass filter depending on the choice of the kernel  $W_{M \times N}$ . That is, if the kernel is set as  $M = N = 1$ , and  $w_{11}=1$ , the LPF( $\cdot$ ) becomes an all-pass filter and, thus,  $d_n = D_n$ .

Next, a point-wise motion detection signal is computed as

$$f_n(i, h) = l_K(d_n(i, h)) \quad (1)$$

where  $l_K(\cdot)$  denotes a linearly scaling function represented as

$$l_K(y) = \begin{cases} 1, & \text{if } y \geq K \\ y/K, & \text{otherwise} \end{cases}$$

in which  $K$  is a positive constant value. Or, as an alternative method, we can define the scaling function  $l_K(y)$  as

$$l_K(y) = \begin{cases} 1, & \text{if } y \geq K \\ (y/K)^a, & \text{otherwise} \end{cases}$$

where  $a$  is a real number. Or, more in general, this invention discloses to use the scaling function  $l_K(y)$  as to be a monotonously increasing function represented as

$$l_K(0)=0, l_K(y)=1 \text{ for } y \geq K, \text{ and}$$

5

$$\forall 0 \leq y_1 \leq y_2 \leq K, l_K(y_1) \leq l_K(y_2)$$

which implies  $l_K(y)$  is monotonously increasing as  $y$  increases. Then the region-wise motion detection signal is computed from the point-wise motion detection signal as

10

$$\phi_n(i, h) = f_n(i, h) + \min(f_{n-1}(i-1, h), f_{n-1}(i+1, h)) \quad (2)$$

where  $f_{n-1}(\cdot)$  denotes the one field delayed motion detection signal described in (1).

Or, as an alternative method, we define the region-wise motion detection signal as

15

$$\phi_n(i, h) = med(f_n(i, h), f_{n-1}(i-1, h), f_{n-1}(i+1, h)) \quad (3)$$

where  $med(\cdot)$  denotes a median operation. Or, in order to minimize the error from a false motion detection, we define the region-wise motion detection signal as

20

$$\phi_n(i, h) = max(f_n(i, h), f_{n-1}(i-1, h), f_{n-1}(i+1, h)) \quad (4)$$

Finally, the region-wise motion detection signal is low pass filtered to form the motion decision parameter  $m_n(i, h)$  as

$$m_n(i, h) = \sum_{p=-a}^b \sum_{q=-c}^d \phi_n(i + 2 \bullet p, h + 2 \bullet q) \cdot a_{p,q} \quad (5)$$

where  $a, b, c, d \geq 0$ , and  $a_{p,q}$  represents a set of normalized (i.e.,  $\sum_{p=-a}^b \sum_{q=-c}^d a_{p,q} = 1$ )

predetermined coefficients of a low pass filter. For instance, the kernel of the low pass filter  
5 used in (5) can be

$$[a_{p,q}'s] = \begin{bmatrix} 0 & 1/8 & 0 \\ 1/8 & 4/8 & 1/8 \\ 0 & 1/8 & 0 \end{bmatrix}$$

The block diagram of Fig. 4 illustrates the computation of the motion decision  
10 parameter  $m_n(i, h)$  as described above with the region-wise motion detection signal defined  
in equation (2).

The block diagram of Fig. 5 illustrates the computation of the motion decision  
parameter  $m_n(i, h)$  as described above with the region-wise motion detection signal defined  
15 in equation (3).

The block diagram of Fig. 6 illustrates the computation of the motion decision  
parameter  $m_n(i, h)$  as described above with the region-wise motion detection signal defined  
in equation (4).

The computed motion decision parameter  $m_n(i,h)$  is then used to mix a spatially interpolated signal and the a temporally interpolated signal, which is described next.

Figs. 7, 8, and 9 are block diagrams of embodiments of the present invention for  
5 interpolating the value of  $x_n(i,h)$  for an interlaced video sequence. The apparatus comprises  
a spatial interpolator 3, a temporal interpolator 4, a motion decision processor 5, and a  
mixer 6. The decision processor 5 corresponds to the diagram illustrated in Figs. 4, 5, and  
6, respectively, and includes, in a signal flow direction, an absolute value former 51 which  
defines the absolute difference parameter  $D_n$ ; a first low pass filter LPF 52 in which the  
10 filtering function  $W_{MxN}$  with the MxN kernel is set; the member 53 represents the linearly  
scaling function as outlined above; a buffer memory 54 and a further line memory 55 are  
connected to a minimum former 56; the function signal  $\phi_n(i,h)$  is formed in an adder 58  
from the signal  $f_n(i,h)$  and the signal from the member 56; finally, the motion detection  
signal  $m_n(i,h)$  is formed by low pass filtering in a second filter, i.e., a spatial low pass filter  
15 LPF 57. The output of the low pass filter 57 is connected so that the motion detection signal  
 $m_n(i,h)$  is supplied to a control input of the mixer 6.

The spatial interpolator 3 spatially interpolates the value of  $x_n(i,h)$  by using a  
predetermined algorithm. The temporal interpolator 4 temporally interpolates the value of  
20  $x_n(i,h)$  by using a predetermined algorithm. The motion decision processor 5 computes  
the motion decision value,  $m_n(i,h)$  as disclosed in the foregoing, which represents the  
degree of the motion at the interpolation location  $(i,h)$ . Conceptually, the value of the  
motion decision parameter will be bounded as  $0 \leq m_n(i,h) \leq 1$  where  $m_n(i,h) = 0$  implies  
“no motion” and  $m_n(i,h) = 1$  implies “motion”. The mixer 6 mixes the output signal

$x_n^s(i,h)$  of the spatial interpolator 3 and the output signal  $x_n'(i,h)$  of the temporal interpolator 4 in accordance with the motion decision value. Where  $x_n^s(i,h)$  and  $x_n'(i,h)$  are the output signal of the spatial interpolator 3 and the output signal of the temporal interpolator 4, respectively, the output signal of the mixer, or, the interpolated signal is  
5 represented as

$$x_n(i,h) = (1 - m_n(i,h)) \cdot x_n'(i,h) + m_n(i,h) \cdot x_n^s(i,h) \quad (6)$$

Note that  $x_n(i,h) = x_n'(i,h)$  when  $m_n(i,h) = 0$  (no motion) and  $x_n(i,h) = x_n^s(i,h)$   
10 when  $m_n(i,h) = 1$  (motion).

In Fig. 7, Fig. 8, and Fig. 9, it does not matter what kind of a spatial interpolating algorithm and a temporal interpolating algorithm are used for the spatial interpolator 3 and the temporal interpolator 4, respectively. The present invention only relates to estimating  
15 the motion decision value  $m_n(i,h)$  and to mixing a spatially interpolated signal and a temporally interpolated signal in accordance with the estimated motion decision value.

Specific information with regard to the interpolation of interlaced video signals and interlaced to progressive conversion is readily available to those of skill in the pertinent art.  
20 The above-noted disclosures, namely U.S. Patent Nos. 5,943,099 and 5,959,681, are specifically incorporated by reference.

10001627-02240

5

Some examples of the spatially interpolated signal  $x_n^s(v, h)$  are

$$x_n^s(i, h) = (x_n(i-1, h) + x_n(i+1, h))/2$$

10

which corresponds to a line average and

$$x_n^s(i, h) = x_n(i-1, h)$$

15

which corresponds to a method known as line doubling.

and

$$x_n^t(i, h) = (x_{n+1}(i, h) + x_{n-1}(i, h))/2$$

$$x_n^t(i, h) = x_{n-1}(i, h)$$